AIR QUALITY ANALYSIS AND PREDICTION IN TAMILNADU

Introduction:-

* Tamil Nadu faces significant challenges related to air quality ,with various regions experiencing high levels of pollution due to industrial activities , vehicular emissions and natural factors.
* Accurate analysis and timely prediction of air quality are essential to mitigate the adverse effects on public health and effect the environment.

Innovative Solutions:-

* Air Quality Monitoring Solutions
* Machine Learning Models
* Satellite Data Integration
* Mobile Apps for Citizen Engagement

Dataset Link:-

<https://tn.data.gov.in/resource/location-wise-daily-ambient-air-quality-tamil-nadu-year-2014>

Columns Details:-

* Stun Code
* State
* City/Down/village
* Location of Monitoring Station
* Agency
* Type of Location
* SO2
* NO2
* RSPM/PM10
* SPM

Libraries Used :

In this we use a libraries like Numpy,Pandas,Matplotlib,Seaborn and import the packages like warnings and filters.

Way to install the libraries:

* Make sure Python and pip is preinstalled on your system.
* To check python version: “python --version”
* To check pip: “pip -V”

Numpy:It is a python library, used to solve numerical problems. It stands for NumericalPython.It can be installed as follows “pip install numpy”.

Matplotlib:Is a python library that helps to plot graphs. It used in data visualizations and graphical plotting. It can be installed as “pip install matplotlib”.

Pandas: Pandas is a python package that is used for data analysis and manipulation. Is a open source libraries that is built over numpy. It can be installed as “pip install pandas”.

Seaborn:Is a python data visualization library based on matplotlib. It can be installed as “pip install seaborn”.

Train and test :

Now we train and test our dataset as follow:

* Data collection: Gather historical air quality data for various locations in Tamil Nadu. This data should include information on pollutants such as PM2.5, PM10, NO2, SO2, CO, O3, and meteorological data like temperature, humidity, wind speed, and direction.
* Data preprocessing:Clean the data by handling missing values, outliers, and duplicates. Perform data normalization or standardization to ensure that all features have similar scales. Create time-based features, such as time of day, day of the week, and season, which can be relevant for air quality prediction.
* Features Selection: Use feature selection techniques to identify the most relevant features for prediction. Correlation analysis and feature importance from machine learning models can help in feature selection.
* Engage local communities in monitoring air quality. Provide them with low-cost air quality monitoring kits and training to collect data in their neighbourhoods. This grassroots approach can supplement official monitoring efforts.
* Launch public awareness campaigns to educate residents about the health risks associated with poor air quality. Promote sustainable practices such as reducing vehicle emissions, planting more trees, and reducing waste burning.

Accuracy:

* Improving the accuracy by splitting the data on heavy variations
* From sklearn import the metrics ad mean square error.
* **Mean squared error (MSE): MSE is a measure of the average squared difference between the predicted and actual values. 4**
* **Mean absolute error (MAE): MAE is a measure of the average absolute difference between the predicted and actual values.**

Conclusion:

Implementing these innovative solutions will require collaboration between government bodies, technology companies, research institutions, and active participation from the community. By combining technological advancements with public engagement, Tamil Nadu can make significant strides in improving its air quality analysis and prediction efforts.